

-continued	
T4 Ligase 4u	1 $\mu$ l
H <sub>2</sub> O	2 $\mu$ l
30	final volume = 10 $\mu$ l

### Transformation of the ligation reaction products into INV $\alpha$ F' competent *E. coli*

### Selection of transformants

Several of the transformants that demonstrated the plasmid with inserts of the correct size were stored in 50% glycerol at  $-80^{\circ}\text{C}$ . Large plasmid DNA stocks were isolated from 250 ml LB overnight cultures for use in COS-7 cell transfections.

65 The eukaryotic cell transfection was carried out according to method described by Chen and Okayama "High-

B

HP  
2/11/03

003060"62E65960

HIGH LEVEL OF EXPRESSION OF INGAP  
IN BACTERIAL AND EUKARYOTIC CELLS

EUKARYOTIC

This is a continuation-in-part application of U.S. Ser. No.  
08/741,096, filed Oct. 30, 1996, now abandoned.

of 08/909, 725 filed August 12, 1997, which is a re-issue application  
of 08/909, 725 filed August 12, 1997, U.S. Patent No.  
5,804,421.

TECHNICAL FIELD OF THE INVENTION  
This invention is related to methods and constructs for  
achieving high level expression of INGAP, a protein  
involved in islet cell neogenesis. 10

BACKGROUND OF THE INVENTION

Pancreatic islets of Langerhans are the only organs of  
insulin production by  $\beta$  cells in the body. However, they 15  
have a limited capacity for regeneration. This limited regen-  
eration capacity predisposes mammals to develop diabetes  
mellitus. Thus there is a need in the art of endocrinology for  
products which can stimulate the regeneration of islets of  
Langerhans to prevent or ameliorate the symptoms of dia- 20  
betes mellitus.

There are many factors regulating pancreatic  $\beta$  cell mass.  
(Vinik, et al., *Diabetes Reviews* 4: 235-263, 1996.) A  
pancreatic extract called ilotropin induces  $\beta$  cell regen- 25  
eration and reverses diabetes. (Rosenberg et al. (1996) *Diabe-*  
*tologia* 39: 256-262. A gene encoding a protein within  
ilotropin has been identified and isolated; the protein is  
responsible for stimulating islet cell regeneration.  
(Rafaeloff, R. *Journal of Clinical Investigations* 99:  
2100-2109, 1997.) This protein is called INGAP, and is 30  
disclosed in patent applications Ser. Nos. 08/401,530,  
08/709,662, and 60/006,271. The disclosure of these appli-  
cations is expressly incorporated herein. Despite the knowl-  
edge of the complete nucleotide sequence of the INGAP 35  
gene, expression of the protein has been limited. Thus there  
is a need in the art for methods of expressing and isolating  
large quantities of the INGAP protein, especially in eukary-  
otic systems.

SUMMARY OF THE INVENTION 40

Sub  
B1 It is an object of the present invention to provide a method  
of producing biologically active INGAP protein from a  
recombinant host cell.

Sub  
B2 It is another object of the present invention to provide a 45  
host cell which expresses large amounts of INGAP protein.

Sub  
B3 It is an object of the present invention to provide a  
recombinant construct for expression of biologically active  
INGAP protein.

Sub  
B4 Another object of the invention is to provide a method for 50  
isolating INGAP protein from a recombinant host cell.

Sub  
B5 These and other objects of the invention are achieved by  
providing the art with a recombinant construct for expres-  
sion of biologically active INGAP protein comprising: 55

a first nucleotide sequence encoding amino acids 27 to  
175 SEQ ID NO: 6 operably linked to a transcriptional  
initiation site and a translational initiation site, wherein a  
second nucleotide sequence encoding a signal peptide is not  
present immediately 5' of said first nucleotide sequence. 60

In another embodiment of the invention a method of  
producing INGAP activity from a recombinant host cell is  
provided. The method comprises the steps of:

culturing a host cell comprising a recombinant construct  
comprising a first nucleotide sequence encoding amino acids 65  
27 to 175 SEQ ID NO: 6 operably linked to a transcriptional  
initiation site and a translational initiation site, wherein a

Sub  
B6 5

In yet another embodiment of the invention a host cell is provided. The host cell comprises a recombinant construct comprising a first nucleotide sequence encoding amino acids 27 to 175 SEQ ID NO: 6 operably linked to a transcriptional initiation site and a translational initiation site, wherein a second nucleotide sequence encoding a signal peptide is not present immediately 5' of said first nucleotide sequence. These and other embodiments of the invention which will be apparent to those of skill in the art provide a practical source of INGAP protein in amounts suitable for use in preclinical and clinical situations.

**FIG. 1. SDS-PAGE gel of products of bacterial transfection.** Bacterial lysate without transfection (CBL), bacterial lysate with transfection (TBL), fractions from Ni-NTA chromatography (eluted at pH6.3 (6.3); pH 5.9 (5.9); and pH 4.5 (4.5) and standards (Std).

25

It is a discovery of the present inventors that bacterial expression as well as eukaryotic expression of INGAP can be achieved at high levels by deleting the coding sequence of the signal sequence of INGAP. While not wanting to be bound by any particular theory or mechanism of action, applicants believe that the signal sequence is toxic to host systems. The signal sequence comprises amino acids 1 to 26 as shown in SEQ ID NO: 5. In the constructions tested, the 5' untranslated region comprising nucleotides 1-16 SEQ ID NO. 1 has also been deleted. This deletion may also contribute to the increase in expression which has been observed.

According to another aspect of the invention, a histidine tag can be put on the protein. The histidine tag can simplify processing and purification. A histidine tag is a stretch of histidine residues which is appended to a protein, usually by genetic engineering. Preferably the tag comprises between 3 and 12 histidine residues. They may be contiguous or interrupted by other residues. The histidine tag may be appended to the N-terminal or to the C-terminal end of the protein to minimize disruption of protein function. Methods for making and utilizing histidine tags are known in the art. The oligohistidine can be used as an affinity moiety using a metal chelate, such as nickel-NTA (N-(5-amino-1-carboxypentyl)-iminodiacetic acid) as the other affinity partner.

60 A recombinant construct according to the invention, is any DNA molecule which has been engineered so that two segments of DNA are adjacent to each other which are not adjacent to each other in nature. Preferably such engineering is performed in vitro, although in vivo engineering can also  
65 be performed. The construct may be a plasmid, phage, virus, transposable element, minichromosome, or other element, as is suitable for the desired application.

In a preferred embodiment, the construct for INGAP expression in eukaryotic systems contains an origin of replication, e.g., EBV oriP, which permits extrachromosomal maintenance in primate and canine cell lines. A sequence element which encodes a nuclear antigen and allows high level episomal replication, e.g., EBNA-1 or SV40 T antigen is also desirable.

*Sub B7* The coding sequence of amino acid residues 27-175 of INGAP protein are included in the constructs. Preferably the entire signal sequence is deleted. However, it is possible that only a portion of the signal sequence must be deleted to obtain excellent expression. Thus some portion of the signal sequence might be retained in the constructs.

Deletion of the 5' untranslated region, nucleotides 1-16, is also desirable. However, it is not known if this is necessary to achieve excellent expression. Thus the 5' untranslated region may be retained in some constructs without departing from the spirit of the invention.

A host cell according to the invention can be transfected or transformed with a recombinant construct according to the present invention. The host may be a bacteria, yeast, insect, or mammalian cell. For eukaryotic expression of INGAP, any cell lines suitable for protein expression may be used, including COS-7 cells and CHO cells.

Selection of suitable promoters and translational initiators for use in the appropriate host cell is well within the ability

of those skilled in the art. For eukaryotic expression system, it is exceedingly useful to choose a promoter sequence which is capable of initiating constitutive transcription to achieve constitutive high level expression of the protein. Rous sarcoma virus long terminal repeat (RSVLTR) is an example of such promoter, although others as are known in the art can be used.

Host cells may be transformed, transfected, mated or infected with the recombinant host cell of the present invention. Culturing of host cells can be performed using techniques and media which are well known in the art. Again, a suitable medium and technique can be selected by an ordinary skilled artisan.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

#### EXAMPLE 1

This example describes the experimental design employed.

We generated a new INGAP cDNA by PCR which excluded the 5' UTR region (nucleotides 1-16 in SEQ ID: 1) and nucleotides encoding the signal peptide (nucleotides 17-94 SEQ ID NO: 1) and created two new restriction enzyme recognition sites enabling the insertion of the new

0090060" 67E9960

construct into a new pQE-31 expression vector. This new ligated construct was transformed into TOP10F<sup>+</sup> competent cells (*E. coli* host strain from Invitrogen). The positive clones were identified, verified by restriction enzyme digestion and the DNA isolated. The DNA was now transformed into a different competent *E. coli* strain, M15(pREP4) and expression of the protein was induced by IPTG (isopropyl-beta, D-thiogalactopyranoside) which inhibits the repressor, facilitating expression of the protein from the M15 promoter/operator. The reason for the intermediate transformation of the ligated material into TOP10F<sup>+</sup> is that these cells are highly competent increasing the odds of getting insert positive colonies. The M15(pREP4) cells that were used for protein expression do not attain competency levels high enough to guarantee transformation of the ligation products. The resultant plasmid DNA obtained from the transformation of the TOP10F<sup>+</sup> was sufficient to enhance the transformation of the M15(pREP4) cells. The His-tagged protein was isolated by Ni<sup>2+</sup> agarose affinity purification.

We used a PCR approach to generate a new INGAP cDNA which excludes the 5' UTR region (nucleotides 1-16 in SEQ ID NO: 1) and nucleotides encoding the signal peptide.

The nucleotide sequence (SEQ ID NO: 1) and corresponding amino acid sequence (SEQ ID NO: 5) that have been excluded are as follows: (the bolded area represents the sequence of the signal peptide)

```

CTGCAAGACAGGTACC    ATG ATG CTT CCC ATG ACC CTC TGT AGG
                        MET MET Leu Pro MET The Leu Cys Arg
ATG TCT TGG ATG CTG CTT TCC TGC CTG ATG TTC CTT TCT TGG
MET Ser Trp MET Leu Leu Ser Cys Leu MET Phe Leu Ser Trp
GTG GAA GGT
Val Glu Gly

```

## EXAMPLE 2

This example describes the use of polymerase chain reaction to synthesize INGMAT (a construct which lacks the signal peptide sequence, i.e., which encodes the mature protein).

### Oligonucleotide design:

Oligonucleotides for PCR were designed to incorporate restriction enzyme recognition sites at their respective 5' ends. The oligonucleotide designed for the 5' end of the gene incorporates a Bam HI site followed by 20 nucleotides corresponding to the N-terminus of the mature protein. The oligonucleotide designed for the 3' end incorporates an Xho I site followed by 20 untranslated nucleotides. The PCR product generated from these primers contains the mature INGAP sequence and the native protein termination codon.

The following is the sequence of the oligonucleotides used:

5' of INGAP (SEQ ID NO: 2)

5'-CCGCGGATCCCGAAGAATCTCAAAAGAAACT-3'

3' of INGAP (SEQ ID NO: 3)

5'-GACCGGCTCGAGTGCTCTTCCTGAGTGAATCC-3'

## 5

### PCR of INGMAT

#### Reaction conditions

Template: (50 ng INGP original cDNA removed from pCDNA3)	5 $\mu$ l	5
MgCl <sub>2</sub> :	4 $\mu$ l	
10 X PCR buffer	5 $\mu$ l	
dATP	1 $\mu$ l	
dCTP	1 $\mu$ l	
dGTP	1 $\mu$ l	10
dTTP	1 $\mu$ l	
5' primer	1 $\mu$ l	
3' primer	1 $\mu$ l	
H <sub>2</sub> O	29 $\mu$ l	
Taq polymerase	1 $\mu$ l	
total volume = 50 $\mu$ l		15

#### cycle parameters

- A) 2 min at 95° C.  
 B) 30 cycles of (1 min 95° C., 1 min 55° C., 1 min 72° C.)  
 C) 7min at 72° C.  
 D) 4° C. until removed from thermal cycler.

The PCR products were then electrophoresed on a 5% PAGE in TBE. Ethidium bromide stained PCR products corresponding to the expected size for the construct were cut from the gel. The gel fragments were electro-eluted into 0.5 ml of TBE, precipitated with 50  $\mu$ l 3M sodium acetate and 1 ml of isopropanol at -80° C. for 20 min, centrifuged, washed once with 1 ml of isopropanol, washed once with 1 ml of 70% ethanol, and then dried under vacuum. The dried pellet was resuspended in 50  $\mu$ l H<sub>2</sub>O and quantified. At the end of this step the sequence of the PCR product that contains both restriction sites minus the signal sequence and 5' UTR was as follows (SEQ ID NO: 4):

```

5'-CC GCG GAT CCC GAA GAA TCT CAA AAG AAA CTGCCT
TCT TCA CGT ATA ACC TGT CCT CAA GGC TCT GTA GCC TAT
GGG TCC TAT TGC TAT TCA CTG ATT TTG ATA CCA CAG ACC
TGG TCT AAT GCA GAA CTA TCC TGC CAG ATG CAT TTC TCA
GGA CAC CTG GCA TTT CTT CTC AGT ACT GGT GAA ATT ACC
TTC GTG TCC TCC CTT GTG AAG AAC AGT TTG ACG GCC TAC
CAG TAC ATC TGG ATT GGA CTC CAT GAT CCC TCA CAT GGT
ACA CTA CCC AAC GGA AGT GGA TGG AGG TGG AGC AGT
TCC AAT GTG CTG ACC TTC TAT AAC TGG GAG AGG AAC CCC
TCT ATT GCT GCT GAC CGT GGT TAT TGT GCA GTT TTG TCT
CAG AAA TCA GGT TTT CAG AAG TGG AGA GAT TTT AAT TGT
GAA AAT GAG CTT CCC TAT ATC TGC AAA TTC AAG GTC TAG
GGC AGT TCT AAT TTC AAC AGC TTG AAA ATA TTA TGA AGC
TCA CAT GGA CAA GGA AGC AAG TAT GAG GAT TCA CTC
AGG AAG AGC ACT CGA GCC GGT C-3'
  
```

♦The bolded areas represent the primers.

## EXAMPLE 3

This example describes the creation of a plasmid containing the expression construct.

Restriction enzyme digestion of the INGMAT PCR product and the pQE-31 vector

We performed two parallel restriction enzyme digestion reactions using 2.5  $\mu$ g of both the INGMAT PCR product and pQE-31 vector. INGMAT was digested with Bam HI and Xho I simultaneously in a 30  $\mu$ l volume. pQE-31 was digested with Bam HI and Sal I simultaneously in a 30  $\mu$ l volume. Both digestion reactions were carried out at 37° C. for a period of 4 hours. After the reactions were completed, 400 ng of each was electrophoresed on a 1.5% agarose gel and stained with ethidium bromide to assure complete digestion. The remainder (~2.1  $\mu$ g) of both digestion reactions were passed over a sepharose G-50 to remove the small DNA fragments followed by two equal volume phenol extractions. The extracted DNA was then precipitated with 2 volumes of ethanol and 1/10 volume 3M sodium acetate at -80° C. for 20 minutes, centrifuged, washed twice with 70% ethanol and dried under vacuum. The pellets were resuspended in 25  $\mu$ l H<sub>2</sub>O and quantified.

The pQE-31 expression system was purchased from QIAGEN Inc, Chatsworth, Calif.

Ligation of INGMAT into pQE-31

INGMAT (Bam HI/Xho I) and pQE-31(Bam HI/Sal I) have compatible ends suitable for ligation. As a result of the ligation the Sal I restriction site in the vector will be eliminated.

Ligation conditions using a 2:1 vector to insert molar ratio.

pQE31(vector) 517 ng	9 $\mu$ l	
INGMAT (insert) 165 ng	2.5 $\mu$ l	
10 X ligation buffer	5 $\mu$ l	
10 mM rATP	5 $\mu$ l	5
T4 Ligase 4u	1 $\mu$ l	
H <sub>2</sub> O	27.5 $\mu$ l	

final volume = 50  $\mu$ l

The ligation reactions were incubated at 4° C. for 16 hours.

#### Transformation of the ligation reaction products into TOP10F<sup>+</sup> competent *E. coli*

We removed 5  $\mu$ l of the ligation reaction into 100  $\mu$ l of competent TOP10F<sup>+</sup> cells, (TOP10F<sup>+</sup> cells were purchased from Invitrogen, San Diego, Calif.) with 0.5  $\mu$ l of 500 mM  $\beta$ -mercaptoethanol and incubated on ice for 30 minutes, heat shocked for 45 seconds at 42° C., and recovered on ice for 2 minutes. Then we added 1 ml of prewarmed S° C. media and incubated at 37° C. with shaking at 225 rpm for 1 hour followed by plating all the transformation reaction on LB broth agar plates containing 100  $\mu$ g/ml ampicillin.

#### Selection of transformants

Colony containing plates were lifted onto Nytran membranes. The colonies were lysed with 0.5M NaOH, neutralized, and the resultant DNA bound to the membrane by baking at 80° C. for 1 hour. The membranes were then hybridized in 50% formaldehyde, 5 $\times$ SSPE at 50° C. for 16 hours with 3 $\times$ 10<sup>6</sup> cpm/ml of <sup>32</sup>P random primed INGAP cDNA. The membranes were washed at high stringency and exposed to X-ray film. Positive colonies were matched up to the X-ray film and grown up in 3 mls of LB with ampicillin.

#### DNA isolation from positive transformants

DNA was isolated from the small cultures using alkaline lysis, phenol extracted, precipitated, dried, and resuspended in 50  $\mu$ l H<sub>2</sub>O. A small aliquot of each of the isolated DNA were digested with Bam HI and Hind III to release inserts. The digested DNAs were electrophoresed on 1.5% agarose and stained with ethidium bromide and positive inserts identified at approximately 510 bp size range. We took four of the insert containing plasmids and incubated them in the presence of RNase to remove any residual bacterial RNA.

#### Transformation of the ligation products into M15 (pREP4) competent *E. coli*

We removed 5  $\mu$ l of the cleaned DNA isolated in section IIE and transfer it into 100  $\mu$ l of M15(pREP4) competent cells. The mixture was incubated on ice for 30 minutes, heat shocked for 45 seconds at 42° C., and recovered on ice for 2 minutes. 1 ml of prewarmed SOC media was added and incubated at 37° C. with shaking at 225 rpm for 90 minutes. All of the transformation reaction was plated on LB broth agar plates containing 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml kanamycin.

#### Selection of transformants for INGMATHIS (INGMAT plus a six-histidine tag) protein production

Eight colonies were picked and grown up in LB with ampicillin. DNA was isolated from the small cultures using alkaline lysis extraction procedures, phenol extracted,



precipitated, dried, and resuspended in 50  $\mu$ l H<sub>2</sub>O. A small aliquot of each of the isolated DNA were digested with Bam HI and Hind III to release inserts. The digested DNA was run on 1.5% agarose gel and visualized by staining with ethidium bromide. Several of the transformants which demonstrated the plasmid with inserts of the correct size as well as the presence of the pREP4 plasmid were stored in 50% glycerol at -80° C. to be used for protein production.

#### EXAMPLE 4

This example describes denaturing metal affinity protein chromatography isolation of the his tagged INGAP protein without signal peptide. (Procedure for a 250 ml pINGMATHIS transformed M15 (pREP4) culture. pINGMATHIS is the INGMATHIS construct ligated into the pQE-31 vector.)

##### Bacteria growth and protein induction

We grew a 25 ml overnight in LB with 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml kanamycin antibiotic. We started a 250 ml LB plus 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml kanamycin culture with 5 ml of the overnight. (1:50) Grown until ABS<sub>600</sub>=0.075 to 0.9 (actual OD=0.866). Added 5 ml of 100 mM IPTG (2 mM final) to induce production of the protein. Continue growing for 4 hours in the case of INGAP. Collected the bacteria and spin at 6000 rpm for 20 minutes, discarded the supernatant. The pellet was frozen until ready to use at -70° C.

##### Ni<sup>+2</sup> NTA agarose preparation

Prepare as much as will be needed. (Use 10 ml of the 50% Ni<sup>+2</sup> NTA for each 250 ml derived bacterial pellet). Place 16 ml of the 50% slurry into a disposable 50 ml centrifuge tube. Centrifuge for 2 minutes at 800xg and discard the supernatant. Add 42 ml of sterile water, resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of sterile water, resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of binding/lysis buffer A (6M Guanidine HCL, 0.1M sodium phosphate, 0.01M Tris, pH 8.0) and resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of binding/lysis buffer A (6M Guanidine HCL, 0.1M sodium phosphate, 0.01M Tris, pH 8.0) and resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Bring the total volume up to 10 ml with buffer A. The slurry is now ready for the application of the lysed bacteria.

##### Bacteria lysis and protein isolation

Thaw the bacterial pellet for 15 minutes at room temperature. Resuspend the pellet in 12.5 ml of lysis buffer A. (6M Guanidine HCL, 0.1M sodium phosphate, 0.01M Tris, pH 8.0). Transfer the resuspension to a 50 ml centrifuge tube. Freeze the resuspension/lysate at -70 until solid. Thaw at room temperature. Place the lysate on a rotator for 60 minutes at room temperature. Centrifuge the lysate for 15 minutes at 10,000xG. Collect the supernatant and add the 10 ml of prepared Ni<sup>2+</sup>NTA. Rotate for 45 minutes. Load the slurry onto a 1.6 cm diameter column and allow to flow through by gravity.

##### Washes

Flow through 50 ml of buffer A. (No need to collect.)  
Flow through 40 ml of buffer B (8M Urea, 0.1M Sodium

000000" 6459960

Wash through 40 ml of buffer E, same as B but pH 4.5. Collect 3 ml fractions. At this point the protein should be in one of the fractions taken. Read the absorbance at 280 of all the fractions to discern where the protein is. Pool, reduce, and SDS page electrophoresis as necessary. 10

In order to purify the expressed protein, we changed the carrier solution of the fraction extracted from the nickel/NTA at pH 4.5 to Tris buffer using dialysis. Dialysis tubing with a molecular weight cut-off of 3000 was prepared by boiling in 5 mM EDTA/200 mM sodium bicarbonate for 5 minutes. The tubing was rinsed briefly in deionized water and boiled another 5 minutes in the bicarbonate solution. The tubing was returned to deionized water, covered with aluminum foil and autoclaved for 10 minutes on a liquid cycle. The tubing was handled with latex gloves during the entire procedure.

One ml of the protein solution from the nickel/NTA column in 6M guanidine HCl was dialyzed against 4 liters of 25 mM Tris buffer at pH 8.5 for 12 hours. After dialysis, there were 2 mls of protein solution with a protein concentration of 800 ug/ml.

This example describes analytical techniques confirming the identity of the product.

Sub  
B8

In order to test for the overexpression of the INGAP protein, discontinuous denaturing polyacrylamide gel electrophoresis was performed on the dialyzed protein solution using the Hoefer SE250 Mighty Small II apparatus. The separating gel was prepared with 15% acrylamide, 1.35% bis-acrylamide in 375 mM Tris buffer at pH 8.8 with 0.05% sodium dodecyl sulfate. Polymerization was induced by addition of 0.05% ammonium persulfate and 20  $\mu$ l TEMED/15 ml solution. The solution was placed in the gel plate apparatus for polymerization. The stacking gel was poured with the same solution, except the Tris buffer was 125 mM at pH 6.8, and the acrylamide concentration was 4%. The protein samples were diluted 1:1 with sample buffer (125 mM Tris-Cl, pH 6.8, 4% SDS, 20% glycerol, and 10% 2-mercaptoethanol).

The upper and lower tank buffers were identical, containing 25 mM Tris, 192 mM glycine and 0.1% SDS at pH 8.3. Two gels were loaded with 20  $\mu$ l each of bacterial lysate without transfection (CBL, 368  $\mu$ g/ml), bacterial lysate with transfection (TBL, 341  $\mu$ g/ml), the fractions from Ni-NTA chromatography (eluted at pH 6.3, 110  $\mu$ g/ml; pH 5.9, 100  $\mu$ g/ml; and pH 4.5, 800  $\mu$ g/ml) and standards (Rainbow Markers, Amersham and Dalton Mark-VII, Sigma). Electrophoresis was performed at 20 mA constant current until the dye front entered the separating gel, and at 60 mA constant current until the dye front reached 0.5 cm from the bottom. The gels were then removed and one was fixed with 45% methanol/10% acetic acid for one hour, and the other

## 5

Sub  
B9 2

2

## 3

3

4

Sub  
BIO

6

Finally, there was a light band at 14 kD recognized by the antibody in both the transfected bacterial lysate and in the pH 4.5 fraction. This may either be another protein or a lytic fraction of the INGAP protein. Given the engineering done to produce the INGAP protein it is most likely a lytic fraction of INGAP.

*Sub B''* In summary, we have been able to express INGAP protein in a prokaryotic system by excluding the 5' UTR and the signal peptide and insertion of the new construct into a new vector. The resultant protein is of the predicted molecular size of INGAP monomer and reacts with the antibody to INGAP in a Western analysis. The protein shares with INGAP peptide the ability to induce ductal cell proliferation.

#### EXAMPLE 6

This example describes the experimental design employed for INGAP expression in eukaryotic systems.

We generated an INGAP cDNA by PCR which excluded the 5' UTR region (nucleotides 1-16 in SEQ ID: 1) and nucleotides encoding the signal peptide (nucleotides 17-94 SEQ ID NO: 1). The reason for excluding the 5' UTR region was to create a protein that is similar to the native protein in which the 5' UTR is not part of the protein. We also created two new restriction enzyme recognition sites enabling the insertion of the new construct into a new pEBVHis-B eukaryotic expression vector. This new ligated construct was transformed into INVαF<sup>+</sup> competent cells (*E. coli* host strain from Invitrogen). The positive clones were identified, verified by restriction enzyme digestion and the DNA isolated and transfected into COS-7 cells. The His-tagged protein was isolated by Ni<sup>2+</sup> agarose affinity purification. The isolated protein showed biological activity when used to stimulate proliferation of ARIP (ductal) cells as measured by <sup>3</sup>H-TdR incorporation.

We used a PCR approach to generate a new INGAP cDNA which excludes the 5' UTR region (nucleotides 1-16 in SEQ ID NO: 1) and nucleotides encoding the signal peptide.

The sequence (SEQ ID NO: 1) that has been excluded is as follows: (the bolded area represents the sequence of the signal peptide)

```
CTGCAAGACAGGTACCATG ATG CTT CCC ATG ACC CTC TGT
      MET MET Leu Pro  MET The  Leu Cys
AGG ATG TCT TGG ATG CTG CTT TCC TGC CTG ATG TTC
Arg  MET Ser  Trp  MET Leu  Leu Ser  Cys  Leu  MET Phe
CTT TCT TGG GTG GAA GGT
Leu Ser  Trp  Val  Glu  Gly
```

To engineer the new INGAP construct we designed oligonucleotides corresponding to the 5' and 3' ends of the INGAP sequence to be amplified.

#### EXAMPLE 7

This example describes the use of polymerase chain reaction to synthesize INGMAT (construct which lacks the signal peptide sequence, i.e., which encodes the mature protein).

##### Oligonucleotide design

Oligonucleotides for PCR were designed to incorporate restriction enzyme recognition sites at their respective 5'

## 12

ends. The oligonucleotide designed for the 5' end of the gene incorporates a Bam HI site followed by 20 nucleotides corresponding to the N-terminus of the mature protein. The oligonucleotide designed for the 3' end incorporates an Xho I site followed by 20 untranslated nucleotides. The PCR product generated from these primers contains the mature INGAP sequence and the native protein termination codon.

The following is the sequence of the oligonucleotides used:

5' of INGAP (SEQ ID NO: 2 )  
 5'-CCGCGGATCCCGAAGAATCTCAAAAGAAACT-3'  
 3' of INGAP (SEQ ID NO: 3)  
 5'- GACCGGCTCGAGTGCTCTTCCTGAGTGAATCC  
 -3'

### PCR of INGMAT

#### Reaction conditions

Template: (50 ng INGAP original cDNA removed from pCDNA3)	5 $\mu$ l
MgCl <sub>2</sub>	4 $\mu$ l
10 X PCR buffer	5 $\mu$ l
dATP	1 $\mu$ l
dCTP	1 $\mu$ l
dGTP	1 $\mu$ l
dTTP	1 $\mu$ l
5' primer	1 $\mu$ l
3' primer	1 $\mu$ l
H <sub>2</sub> O	29 $\mu$ l
Taq polymerase	1 $\mu$ l
total volume = 50 $\mu$ l	

#### cycle parameters

- A) 2 min at 95° C.
- B) 30 cycles of (1 min 95° C., 1 min 55° C., 1 min 72° C.)
- C) 7 min at 72° C.
- D) 4° C. until removed from thermal cycler.

The PCR products were then electrophoresed on a 5% PAGE in TBE. Ethidium bromide stained PCR products corresponding to the expected size for the construct were cut from the gel. The gel fragments were electro-eluted into 0.5 ml of TBE, precipitated with 50  $\mu$ l 3M sodium acetate and 1 ml of isopropanol at -80° C. for 20 min, centrifuged, washed once with 1 ml of isopropanol, washed once with 1 ml of 70% ethanol, and then dried under vacuum. The dried pellet was resuspended in 50  $\mu$ l H<sub>2</sub>O and quantified. At the end of this step the sequence of the PCR product that contains both restriction sites minus the signal sequence and 5' UTR was as follows (SEQ ID NO: 4):

5'-CC GCG GAT CCC GAA GAA TCT CAA AAG AAA CTGCCT  
TCT TCA CGT ATA ACC TGT CCT CAA GGC TCT GTA GCC TAT  
GGG TCC TAT TGC TAT TCA CTG ATT TTG ATA CCA CAG ACC  
TGG TCT AAT GCA GAA CTA TCC TGC CAG ATG CAT TTC TCA  
GGA CAC CTG GCA TTT CTT CTC AGT ACT GGT GAA ATT ACC  
TTC GTG TCC TCC CTT GTG AAG AAC AGT TTG ACG GCC TAC  
CAG TAC ATC TGG ATT GGA CTC CAT GAT CCC TCA CAT GGT  
ACA CTA CCC AAC GGA AGT GGA TGG AGG TGG AGC AGT  
TCC AAT GTG CTG ACC TTC TAT AAC TGG GAG AGG AAC CCC  
TCT ATT GCT GCT GAC CGT GGT TAT TGT GCA GTT TTG TCT  
CAG AAA TCA GGT TTT CAG AAG TGG AGA GAT TTT AAT TGT  
GAA AAT GAG CTT CCC TAT ATC TGC AAA TTC AAG GTC TAG  
GGC AGT TCT AAT TTC AAC AGC TTG AAA ATA TTA TGA AGC  
TCA CAT GGA CAA GGA AGC AAG TAT GAG GAT TCA CTC  
AGG AAG AGC ACT CGA GCC GGT C-3'

◆The bolded areas represent the primers.

#### EXAMPLE 8

This example describes the creation of a plasmid containing an expression construct for expression in eukaryotic systems.

##### Restriction enzyme digestion of the INGMAT PCR product and the pEBVHis-B vector

We performed two parallel restriction enzyme digestion reactions using 2.5  $\mu$ g of both the INGMAT PCR product and pEBVHis-B vector. INGMAT was digested with Bam HI and Xho I simultaneously in a 30  $\mu$ l volume. pEBVHis-B was digested with Bam HI and Xho I simultaneously in a 30  $\mu$ l volume. Both digestion reactions were carried out at 37° C. for a period of 4 hours. After the reactions were completed, 400 ng of each was electrophoresed on a 1.5% agarose gel and stained with ethidium bromide to assure complete digestion. The remainder (~2.1  $\mu$ g) of both digestion reactions were passed over a sepharose G-50 column twice to remove the small DNA fragments followed by two equal volume phenol extractions. The extracted DNA was then precipitated with 2 volumes of ethanol and 1/10 volume 3M sodium acetate at -80° C. for 20 minutes, centrifuged, washed twice with 70% ethanol and dried under vacuum. The pellets were resuspended in 25  $\mu$ l H<sub>2</sub>O and quantified.

The pEBVHis-B expression system was purchased from INVITROGEN Corp. San Diego, Calif.

##### Ligation of INGMAT into pEBVHis-B

INGMAT (Bam HI/Xho I) and pEBVHis-B(Bam HI/Xho I) have compatible ends suitable for ligation.

Ligation conditions using a 20:1 insert to vector molar ratio.

pEBVHis-B(vector) 62 ng	1 $\mu$ l	
INGMAT (insert) 80 ng	4 $\mu$ l	
10 X ligation buffer	1 $\mu$ l	65
10 mM rATP	1 $\mu$ l	

COS-7 cells (SV40 transformed African green monkey kidney cells) were grown on twenty 150 mm diameter plates in culture medium (Dulbecco's modified Eagles medium, 10% fetal bovine serum, penicillin/streptomycin) to 80% confluency.

Each plate was washed twice with 10 mls of PBS, and 2' mls of fresh culture medium added. The DNA transfection mixture (2.5 ml) was added dropwise to each plate, swirled gently, and incubated overnight at 37° C.

DNA transfection mixture:

pEBVHis-INGMAT 60 ng	0.080 ml
H <sub>2</sub> O	1.045 ml
2.5 M CaCl <sub>2</sub>	0.125 ml
2XBES	1.25 ml
<hr/>	
Final volume =	2.5 ml

The transfection media was removed from the plates. The plates were then washed 3 times with culture medium, replenished with 25 ml of culture media, and incubated for 48 hrs. The plates were washed twice with PBS and trypsinized. The trypsinized cells were collected from groups of 5 plates, pelleted, and frozen with liquid nitrogen.

Sub  
B'2 ch

~~This example describes denaturing metal affinity protein chromatography isolation of his tagged-INGAP protein without signal peptide. (Procedure for 2 cell pellets from five 150 mm plates each of pEBVHis-INGMAT transfected COS-7 cells)~~

### Ni<sup>2+</sup> NTA agarose preparation

Place 5 ml of the 50% slurry into a disposable 50 ml centrifuge tube. Centrifuge for 2 minutes at 800xg and discard the supernatant. Add 42 ml of sterile water, resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of sterile water, resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of binding/lysis buffer A (6M Guanidine HCl, 0.1M sodium phosphate, 0.01M Tris, pH 8.0) and resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml of binding/lysis buffer A (6M Guanidine HCL, 0.1M sodium phosphate, 0.01M Tris, pH 8.0) and resuspend the resin. Centrifuge for 2 minutes at 800xG and discard the supernatant. Add 42 ml

[illegible]

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 94 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear



( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:1:

( 2 ) INFORMATION FOR SEQ ID NO:2:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 31 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCGCGGATCC CGAAGAATCT CAAAAGAAAC T 31

( 2 ) INFORMATION FOR SEQ ID NO:3:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 32 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

· ( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACCGGCTCG AGTGCTCTTC CTGAGTGAAT CC 32

( 2 ) INFORMATION FOR SEQ ID NO:4:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 558 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: cDNA

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCGCGGATCC	CGAAGAATCT	CAAAAAGAAAC	TGCCTTCTTC	ACGTATAACC	TGTCCTCAAG	60
GCTCTGTAGC	CTATGGGTCC	TATTGCTATT	CACTGATTTT	GATACCACAG	ACCTGGTCTA	120
ATGCAGAACT	ATCTGCCAG	ATGCATTTCT	CAGGACACCT	GGCATTTCCT	CTCAGTACTG	180
GTGAAATTAC	CTTCGTGTCC	TCCCTTGTGA	AGAACAGTTT	GACGGCCTAC	CAGTACATCT	240
GGATTGGACT	CCATGATCCC	TCACATGGTA	CACTACCCAA	CGGAAGTGGA	TGGAAGTGGA	300
GCAGTTCCAA	TGTGCTGACC	TTCTATAACT	GGGAGAGGAA	CCCCTCTATT	GCTGCTGACC	360
GTGGTTATTG	TGCAGTTTTG	TCTCAGAAAT	CAGGTTTTCA	GAAGTGGAGA	GATTTTAATT	420
GTGAAAATGA	GCTTCCCTAT	ATCTGCAAAT	TCAAGGTCTA	GGGCAGTTCT	AATTTCAACA	480
GCTTGAAAAT	ATTATGAAGC	TCACATGGAC	AAGGAAGCAA	GTATGAGGAT	TCACTCAGGA	540
AGAGCACTCG	AGCCGGTC					558

( 2 ) INFORMATION FOR SEQ ID NO:5:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 26 amino acids  
( B ) TYPE: amino acid  
( C ) STRANDEDNESS: single  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: peptide

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Met Leu Pro Met Thr Leu Cys Arg Met Ser Trp Met Leu Leu Ser  
1 5 10 15

Cys Leu Met Phe Leu Ser Trp Val Glu Gly  
20 25

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 175 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: linear

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met 1	Met	Leu	Pro	Met 5	Thr	Leu	Cys	Arg	Met 10	Ser	Trp	Met	Leu	Leu 15	Ser
Cys	Leu	Met	Phe 20	Leu	Ser	Trp	Val	Glu 25	Gly	Glu	Glu	Ser	Gln 30	Lys	Lys
Leu	Pro	Ser 35	Ser	Arg	Ile	Thr	Cys 40	Pro	Gln	Gly	Ser	Val 45	Ala	Tyr	Gly
Ser	Tyr 50	Cys	Tyr	Ser	Leu	Ile 55	Leu	Ile	Pro	Gln	Thr 60	Trp	Ser	Asn	Ala
Glu 65	Leu	Ser	Cys	Gln	Met 70	His	Phe	Ser	Gly	His 75	Leu	Ala	Phe	Leu	Leu 80
Ser	Thr	Gly	Glu	Ile 85	Thr	Phe	Val	Ser	Ser 90	Leu	Val	Lys	Asn 95	Ser	Leu
Thr	Ala	Tyr	Gln 100	Tyr	Ile	Trp	Ile	Gly 105	Leu	His	Asp	Pro	Ser 110	His	Gly
Thr	Leu	Pro 115	Asn	Gly	Ser	Gly	Trp 120	Lys	Trp	Ser	Ser	Ser 125	Asn	Val	Leu
Thr	Phe 130	Tyr	Asn	Trp	Glu	Arg 135	Asn	Pro	Ser	Ile	Ala 140	Ala	Asp	Arg	Gly
Tyr 145	Cys	Ala	Val	Leu	Ser 150	Gln	Lys	Ser	Gly	Phe 155	Gln	Lys	Trp	Arg	Asp 160
Phe	Asn	Cys	Glu	Asn 165	Glu	Leu	Pro	Tyr	Ile 170	Cys	Lys	Phe	Lys	Val 175	

We claim:

- We claim:
1. A recombinant construct for expression of Islet Neo-45 genesis Associated Protein or INGAP activity comprising:
    - a first nucleotide sequence encoding amino acids 27 to 175 as shown in SEQ ID NO: 6 operably linked to a transcriptional initiation site and a translational initiation site, wherein a second nucleotide sequence encoding a signal peptide is not present immediately 5' of said first nucleotide sequence. 50
  2. The construct of claim 1 wherein nucleotides 1-16 of SEQ ID NO: 1 are not present 5' of said first nucleotide sequence.
  3. The construct of claim 1 further comprising a third nucleotide sequence encoding a histidine tag. 55
  4. The construct of claim 3 wherein the third nucleotide sequence is immediately 5' or 3' to said first nucleotide sequence.
  5. The construct of claim 1 wherein the transcriptional 60 initiation site is inducible.
  6. The construct of claim 1 wherein the transcriptional initiation site is the lac promoter/operator.
  7. The construct of claim 1 further comprising a promoter sequence capable of initiating constitutive transcription.
  8. The construct of claim 7 wherein the promoter 65 sequence is Rous sarcoma virus long terminal repeat (RSVLTR).

45 9. The construct of claim 1 further comprising a nucleotide sequence encoding a nuclear antigen.

10. The construct of claim 9 wherein the nuclear antigen is EBNA-1.

11. The construct of claim 1 further comprising an origin of replication.

50 12. The construct of claim 11 wherein the origin of replication is Epstein Bar Virus (EBV) origin of replication.

13. A method of producing biologically active Islet Neogenesis Associated Protein or INGAP protein from a recombinant host cell comprising the steps of:

55 culturing a host cell comprising a recombinant construct comprising a first nucleotide sequence encoding amino acids 27 to 175 as shown in SEQ ID NO: 6 operably linked to a transcriptional initiation site and a translational initiation site, wherein a second nucleotide sequence encoding a signal peptide is not present immediately 5' of said first nucleotide sequence, and  
60 recovering protein from said cultured host cell.

14. The method of claim 13 wherein the construct further comprises a third nucleotide sequence encoding a histidine tag, and INGAP protein is purified using a nickel affinity  
65 matrix.

15. A host cell comprising a recombinant construct comprising a first nucleotide sequence encoding amino acids 27

chsbw

chsbw

000000-000000



17. The method of claim 13 wherein the first nucleotide sequence encoding amino acids 27-175 comprises nucleotides 12-456 of SEQ ID NO: 4.

18. The host cell of claim 15 wherein the first nucleotide sequence encoding amino acids 27-175 comprises nucleotides 12-456 of SEQ ID NO: 4.

\* \* \* \* \*

008060" 62E65960